


esiCancer: Evolutionary *In Silico* Cancer Simulator

Darlan Conterno Minussi^{1,2}, Bernardo Henz³, Mariana dos Santos Oliveira^{1,2},
Eduardo C. Filippi-Chiela⁴, Manuel M. Oliveira³, and Guido Lenz^{1,2}



Abstract

The evolution of cancer is inferred mainly from samples taken at discrete points that represent glimpses of the complete process. In this study, we present esiCancer as a cancer-evolution simulator. It uses a branching process, randomly applying events to a diploid oncogenome, altering probabilities of proliferation and death of the affected cells. Multiple events that occur over hundreds of generations may lead to a gradual change in cell fitness and the establishment of a fast-growing population. esiCancer provides a platform to study the impact of several factors on tumor evolution, including dominance,

fitness, event rate, and interactions among genes as well as factors affecting the tumor microenvironment. The output of esiCancer can be used to reconstruct clonal composition and Kaplan-Meier-like survival curves of multiple evolutionary stories. esiCancer is an open-source, standalone software to model evolutionary aspects of cancer biology. 

Significance: This study provides a customizable and hands-on simulation tool to model the effect of diverse types of genomic alterations on the fate of tumor cells.

Introduction

Despite immense advances, the study of the molecular biology of cancer (1–3) remains dependent on biopsies, restricted to specific timepoints and to a fraction of the whole tumor. Thus, they fail to capture a complete picture of cancer heterogeneity offering only a snapshot of tumor evolution. The evolutionary story from a normal cell to a heterogeneous population of billions of cells is complex and, therefore, requires new theoretical insights to better understand the process.

Several models have been developed to study cancer *in silico* (4), each focusing on a specific characteristic of cancer biology. These include the hallmarks of cancer (5), the rate of clonal expansion (6), stem-cell-driven tumor initiation (7), the effect of cell migration on tumor growth (8), and the impact of the microenvironment on tumor evolution (9). With esiCancer, we provide a fully customizable tool designed to help stitch together genetic events during cancer clonal evolution. esiCancer follows a stochastic branching model. Its simulations generate evolutionary paths with events that modify the fitness of cells, leading to the selection of the fittest ones.

¹Departamento de Biofísica, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Rio Grande do Sul, Brazil. ²Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Rio Grande do Sul, Brazil. ³Instituto de Informática, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Rio Grande do Sul, Brazil. ⁴Departamento de Ciências Morfológicas, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Rio Grande do Sul, Brazil.

Note: Supplementary data for this article are available at Cancer Research Online (<http://cancerres.aacrjournals.org/>).

D.C. Minussi, B. Henz, and M.S. Oliveira contributed equally to this article.

Corresponding Author: Guido Lenz, Universidade Federal do Rio Grande do Sul (UFRGS), Av. Bento Gonçalves, n. 9500, Predio 43431, Sala 107, Porto Alegre, Rio Grande do Sul 91501-970, Brazil. Phone: 5551-3308-7620; Fax: 5551-3308-7003; E-mail: lenz@ufrgs.br

doi: 10.1158/0008-5472.CAN-17-3924

©2018 American Association for Cancer Research.

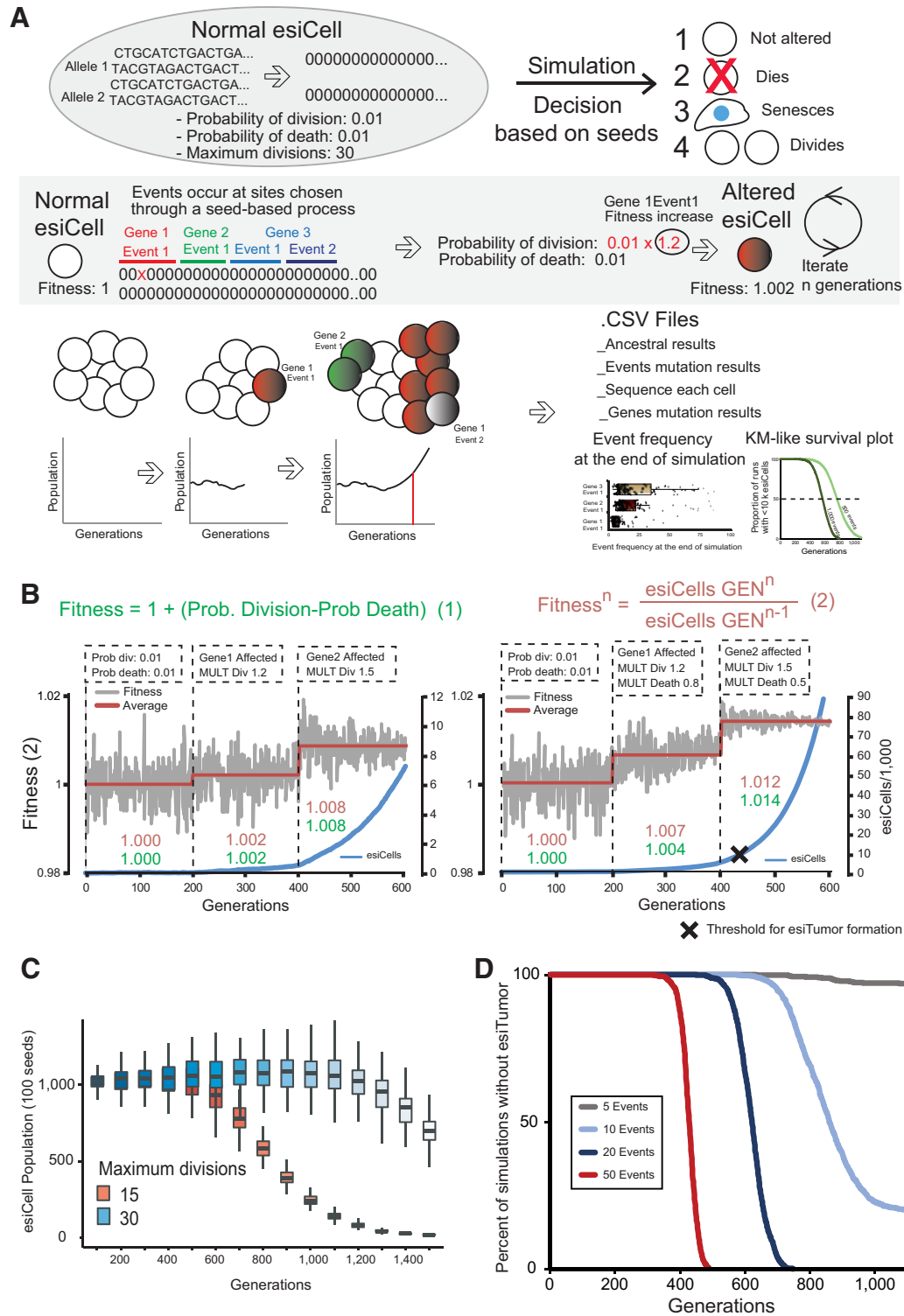
Materials and Methods

esiCancer

esiCancer simulates a population of esiCells, each containing a diploid representation of its genome as two independent lists, a probability of death, a probability of division, and a maximum number of divisions (Fig. 1A). This genome can be hit by genetic events, representing point mutations, translocations, indels, etc., with a defined probability and dominance. These events alter the fitness and other aspects of the affected esiCell (Supplementary Video S1).

esiCancer applies events to a predefined number of esiCells, each one independently subjected to four possible outcomes: no alteration; death; senescence; or cell division (Fig. 1A). If an esiCell divides, the two daughter cells receive, at random sites, a number of genetic events, defined by the user. Each event is associated with a change in the probability of division, death, mutation, and/or maximum divisions, thus impacting the population of esiCells over time. For all stochastic decisions, esiCancer uses a pseudorandom number generator initialized with a seed value. Different seeds create different evolutionary stories, which can be automatically iterated over multiple seeds to grant high-throughput simulations. A given seed will re-create the same sequence of events thus guaranteeing reproducibility (Supplementary Video S2). esiCancer exports data about the cell lineages, the sequence, and frequency of events that gave rise to specific groups of esiCells, providing a complete analysis of the clonal composition of an esiTumor (Fig. 1A; Supplementary Video S3).

Precompiled Linux, Windows, and MacOS GUI-based versions of esiCancer, as well as examples of esiTables, outputs, and video tutorials outlining how to use the system and analyze its output data are available at <http://www.ufrgs.br/labsinal/esiCancer/>. There one can also find detailed documentation about esiCancer, which includes pipelines to assist users in selecting the oncogene and the parameters for their simulations. A guide for the production of the figures presented in this report is also provided. Source code and additional information can be found at <https://github.com/bernardohenz/esiCancer>. esiCancer is under GNU Public License v3.0.

**Figure 1.**

Overview of esiCancer. **A**, Modeling starts with normal esiCells predefined probabilities of division and death, and a number of maximum divisions; if an esiCell divides, it receives a number of events and if an event hits a gene, characteristics defined in the esiTable are applied. Number of esiCells, their mutations, and fitness are recorded. **B**, Stochasticity of fitness of the esiCell population. Fitness in esiCancer is defined by the number of esiCells in generation n divided by the number of cells at generation $n-1$ (1) or by the difference between the probability of cell division and death (2). Fitness using formula 1 in green or 2 in red. GEN, generation. **C**, Population of esiCells with 15 or 30 maximum divisions for 100 different seeds. **D**, Kaplan-Meier (KM)-like plots for runs with different event rates. For these runs, 10k esiCells from an initial population of 1,000 esiCells were considered an esiTumor.

Randomness of the population fitness

Fitness in evolutionary biology is defined by the number of individuals in the n^{th} generation (GEN^n) divided by the number of individuals in the previous generation (GEN^{n-1}). In esiCancer, fitness is directly defined by the probability of division minus the probability of death (Fig. 1B). If the probability of division and death are both set to 0.01, fitness value calculated with the input data (Fig. 1B, equation 1) is similar to the value calculated with the output values (Fig. 1B, equation 2), and this continues to be true after alteration in fitness produced by events. An event that affects the probability of division increases the average fitness, which is further increased by a second event. If an event increases the probability of division and decreases the probability of death, the impact of this event on the fitness reflects the impact on both division and death (Fig. 1B). As expected with exponential growth, this produces a final number of esiCells, which is about 8 times higher when compared with the impact of only increasing the probability of division.

Escape from replicative senescence is another important hallmark of cancer. esiCancer allows the user to limit the number of divisions, resulting in a gradual reduction in the population because cells retain their probability of death (Fig. 1C). Events that lead to an increase in the maximum number of divisions model an escape from replicative senescence. esiCancer can also be used to generate Kaplan-Meier-like graphs by plotting the number of generations required to achieve a defined threshold. Increasing the number of events per division also increases the

number of simulations that reach the threshold while reducing the number of generations required to reach such condition (Fig. 1D).

Survival of the fittest

In esiCancer, different simulations produce unique frequencies in gene events, but the frequency after 1,100 generations of a given event on average directly correlates with its dominance (Fig. 2A, i), probability (Fig. 2A, iii), and impact on the fitness (Fig. 2A, ii) as predicted by evolutionary biology. Highly dominant events will appear more frequently than events with low dominance, as the impact of a mutation on the first allele of a highly dominant gene is much stronger than on genes with low dominance values (Fig. 2A, i). Gene frequency also directly correlates with fitness (Fig. 2A, ii) and the probability of the event (Fig. 2A, iii). Therefore, these parameters will affect the probability of an event occurring and will alter the number of descendants that contain the event. A given gene can have two events, which interact allelically and, if all other conditions are the same, their frequency is higher than the frequency of a gene with a single event (Fig. 2A, iv).

An event can also impact several genes, resembling copy number variation (CNV). An event affecting gene A and B, but not C, will have a frequency equal to gene A, if gene A does not receive any additional event by itself. Frequency of gene B will be the sum of the frequencies due to event AB and an additional event on gene B. Event C will not be affected by event AB (Fig. 2A, right). Finally, the relative frequency of events at different timepoints indicates that the same conditions, when modeled with different

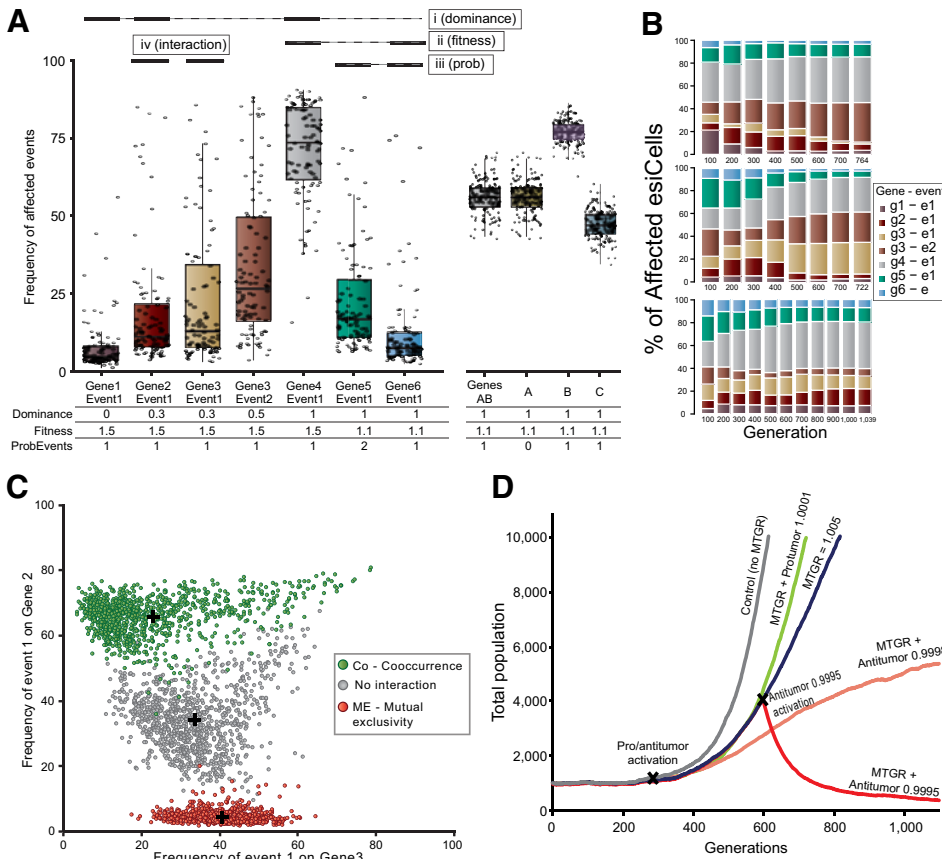


Figure 2. **A**, Frequency of events with the indicated dominance, fitness, interaction, and probability after 1,100 generations, with 20 events per generation. Each point represents a simulation with a different seed. Median \pm quartiles. Numbers indicate the comparisons mentioned in the text. **B**, Dynamics in the frequency of events during esiCancer simulation. **C**, Impact of ME or CO between gene 1 and 2 on the frequency of events in gene 2 and 3. **D**, Impact of the environment on the total population of a defined seed. Maximal tumor growth rate (MTGR) can be set to restrict tumor growth, which can be impacted by events that increase or decrease this value.

Downloaded from http://aacrjournals.org/cancerres/article-pdf/79/5/1012/2789293/1010.pdf by guest on 08 October 2024

seeds, can produce variable population dynamics recapitulating different models of tumorigenesis (Fig. 2B).

Gene and cell interactions in esiCancer

Cancer genes act within complex interaction networks during tumor development. A given event can affect the impact of another event, either by decreasing its impact, leading to mutual exclusivity, or increasing its impact, resulting in cooccurrence (10). In a simulation containing 3 genes with equal settings and no interactions, a similar frequency of event 1 in gene 2 and 3 occurs (Fig. 2C, gray). If the impact of gene 1 and 2 are mutually exclusive, gene 2 will appear less frequently altered when compared with noninteracting genes and the contrary occurs in the case of cooccurrence (Fig. 2B, red and green). esiCancer also permits the modeling of interactions among cells, in which events can have impacts on the whole tumor, resulting in alterations that impact the microenvironment positively or negatively (Fig. 2D).

Conclusion

esiCancer provides a platform for simulating the genetics of tumor evolution. It was designed from the ground up to model important aspects of evolutionary biology applied to cancer using real genetic data. The unique strategy of modeling individual cells and applying single-cell decisions of division, senescence, or death reproduces key aspects of tumorigenesis. This results in the survival of the fittest, where each simulation yields a unique

outcome, thereby resembling the rise of cancer in humans and capable of modeling the response to mutagens or genetic alterations. In this way, esiCancer can become an important tool to better understand the hidden aspects of tumor evolution.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: D.C. Minussi, B. Henz, M.M. Oliveira, G. Lenz
Development of methodology: D.C. Minussi, B. Henz, M. Oliveira, E.C. Filippi-Chiela, G. Lenz

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): D.C. Minussi, M. Oliveira, E.C. Filippi-Chiela, G. Lenz

Writing, review, and/or revision of the manuscript: D.C. Minussi, B. Henz, M. Oliveira, E.C. Filippi-Chiela, M.M. Oliveira, G. Lenz

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases):

Study supervision: G. Lenz

Others (supervision of computer program development): M.M. Oliveira

Acknowledgments

This work was supported by FAPERGS/PRONEX (16-2551). All authors are or were recipients of fellowships from CNPq. We wish to thank Dr. Francisco M. Salzano (in memoriam) and Francisco Ivanio for critical reading of the manuscript and Maria Julia Oliveira for video and sound editing.

Received December 29, 2017; revised September 18, 2018; accepted December 14, 2018; published first December 18, 2018.

References

- Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, et al. Mutational landscape and significance across 12 major cancer types. *Nature* 2013;502:333–9.
- Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, Golub TR, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* 2014;505:495–501.
- McGranahan N, Swanton C. Clonal heterogeneity and tumor evolution: past, present, and the future. *Cell* 2017;168:613–28.
- Beerenwinkel N, Schwarz RF, Gerstung M, Markowitz F. Cancer evolution: mathematical models and computational inference. *Syst Biol* 2015;64:e1–25.
- Spencer SL, Gerety RA, Pienta KJ, Forrest S. Modeling somatic evolution in tumorigenesis. *PLoS Comput Biol* 2006;2:e108.
- Bozic I, Antal T, Ohtsuki H, Carter H, Kim D, Chen S, et al. Accumulation of driver and passenger mutations during tumor progression. *Proc Natl Acad Sci U S A* 2010;107:18545–50.
- Gentry SN, Jackson TL. A mathematical model of cancer stem cell driven tumor initiation: implications of niche size and loss of homeostatic regulatory mechanisms. *PLoS One* 2013;8:e71128.
- Waclaw B, Bozic I, Pittman ME, Hruban RH, Vogelstein B, Nowak MA. A spatial model predicts that dispersal and cell turnover limit intratumour heterogeneity. *Nature* 2015;525:261–4.
- Lloyd M, Cunniff JJ, Bui MM, Gilles RJ, Brown JS, Gatenby RADarwinian dynamics of intratumoral heterogeneity: not solely random mutations but also variable environmental selection forces. *Cancer Res* 2016;76:3136–44.
- Mina M, Raynaud F, Tavernari D, Battistello E, Sungalee S, Saghafinia S, et al. Conditional selection of genomic alterations dictates cancer evolution and oncogenic dependencies. *Cancer Cell* 2017;32:155–68.